REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY): 14-05-2014

2. REPORT TYPE: Final Report

3. DATES COVERED (From - To): 15-Oct-2013 - 14-Apr-2014

4. TITLE AND SUBTITLE: Construction of 3-D Terrain Models from BIG Data Sets

5a. CONTRACT NUMBER

5b. GRANT NUMBER: W911NF-13-P-0018

5c. PROGRAM ELEMENT NUMBER: 665502

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Pankaj K. Agarwal, Thomas Moelhave

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Scalable Algorithmics USA
2400 Huntscroft Ln
Apt 203
Raleigh, NC 27617 -8503

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S): ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S): 64287-CS-ST1.3

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Report developed under Topic #A13A-T005, contract W911NF-13-P-0018.

The objective of this Phase I project was to investigate the feasibility of scalable, efficient and practical algorithms for analysis-driven construction of high-resolution 3D terrain models from BIG terrain data sets, and to build a prototype software

15. SUBJECT TERMS
gis,Delaunay triangulation,visibility,data uncertainty, denoising, contour tree, topology,STTR report

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Pankaj Agarwal
					19b. TELEPHONE NUMBER: 919-660-6540

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Construction of 3-D Terrain Models from BIG Data Sets

ABSTRACT

Report developed under Topic #A13A-T005, contract W911NF-13-P-0018.

The objective of this Phase I project was to investigate the feasibility of scalable, efficient and practical algorithms for analysis-driven construction of high-resolution 3D terrain models from BIG terrain data sets, and to build a prototype software infrastructure for making analysis-prepared terrain models available to data consumers on multiple platforms. Analysis-driven modeling means that the construction of the model is influenced by, and prepared for, the specific analysis that the terrain model will be used for by data consumers.

The following list summarizes the main findings of Phase I:

* Constructing and maintaining DEM. A practical algorithm and its implementation for constructing large (static) constrained Delaunay triangulations
* Denoising. A proof-of-concept bridge-detection algorithm that can detect likely locations of bridges having a significant impact on flow networks.
* Analysis-driven level of details. An algorithmic framework for building a visibility-preserving hierarchical representation of a terrain.
* Uncertainty-aware algorithms. An efficient algorithm for constructing a stochastic model of a terrain that incorporates uncertainty in LiDAR data.
* Online service. A simple prototype of an online service.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received		Paper
05/14/2014	2.00	D. Shaharabani, O. Salzman, P. K. Agarwal, D. Halperin. Sparsification of motion-planning roadmaps by edge contraction, International Journal of Robotics Research (03 2014)

TOTAL: 1

Number of Manuscripts:

Books

Received	Paper

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Wuzhou Zhang	0.74	
You Wu	0.33	
FTE Equivalent:	**1.07**	
Total Number:	**2**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Pankaj K. Agarwal	0.08	
FTE Equivalent:	**0.08**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	Discipline
Niel Lebeck	0.00	Computer Science
FTE Equivalent:	**0.00**	
Total Number:	**1**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 1.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 1.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 1.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 1.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 1.00

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME
Total Number:

Names of other research staff

NAME	PERCENT_SUPPORTED
Thomas Moelhave	0.80
FTE Equivalent:	**0.80**
Total Number:	**1**

Sub Contractors (DD882)

1 a. Duke University

1 b. c/o Office of Research Support
2200 W Main St Ste 710
Durham NC 277054677

Sub Contractor Numbers (c):

Patent Clause Number (d-1):

Patent Date (d-2):

Work Description (e):

Sub Contract Award Date (f-1): 10/15/13 12:00AM

Sub Contract Est Completion Date(f-2): 4/14/14 12:00AM

1 a. Duke University

1 b. Office of Sponsored Programs
Box 90491
Durham NC 277080491

Sub Contractor Numbers (c):

Patent Clause Number (d-1):

Patent Date (d-2):

Work Description (e):

Sub Contract Award Date (f-1): 10/15/13 12:00AM

Sub Contract Est Completion Date(f-2): 4/14/14 12:00AM

Inventions (DD882)

Scientific Progress

Technology Transfer

Construction of 3-D Terrain Models from BIG Data Sets

Scalable Algorithmics USA (SCALGO USA)

1 Technical Objectives of Phase I

The objective of Phase I was to demonstrate the feasibility of constructing analysis-prepared digital elevation models (APDEMs) from big heterogeneous data, and of constructing a software infrastructure for making these models available to data consumers. To this end, Phase I focused on designing algorithms for a number of problems — and implementing a subset, which also led to a basic prototype of the full system. Some of these problems (e.g. constrained Delaunay triangulation; denoising) built on our past work, while others were relatively unexplored (e.g. uncertainty-aware algorithms; analysis-driven levels of detail). More specifically, we divided the goal of preparing APDEMs into five tasks and proposed to study the following specific problems for each of the tasks in Phase I:

(C1) Constructing and maintaining DEM. A practical algorithm and its implementation for constructing large (static) constrained Delaunay triangulations with large amounts of constraints.

(C2) Denoising. A proof-of-concept bridge-detection algorithm that can detect likely locations of bridges having a significant impact on flow networks.

(C3) Analysis-driven level of details. An algorithmic framework for building a visibility-preserving hierarchical representation of a terrain.

(C4) Uncertainty-aware algorithms. An efficient algorithm for constructing a stochastic model of a terrain that incorporates uncertainty in LiDAR data.

(C5) Online service. A simple prototype of an online service that supports: (i) simple DEM retrievals, (ii) submissions of data corrections and updates from users, and (ii) visualization of the updated results.

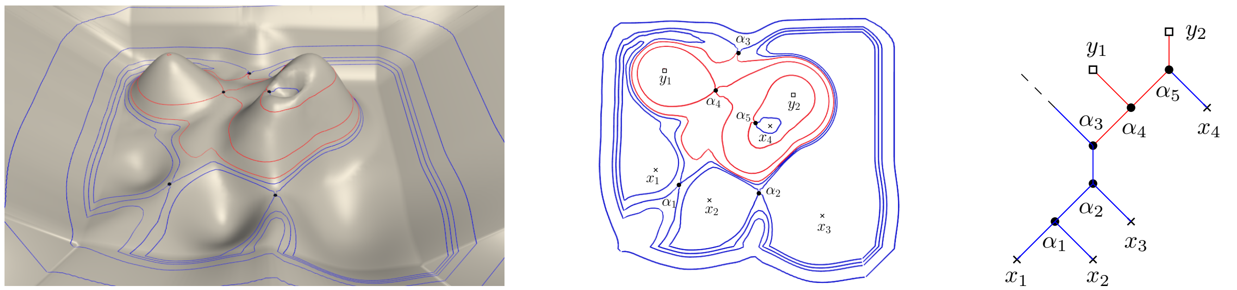
1.1 Phase I findings

This subsection summarizes the research performed in Phase I, and the main accomplishments of this phase.

C1 - Constructing and maintaining APDEMs. *Constrained Delaunay triangulation.* The Phase I objective was to show the feasibility to construct nationwide APDEMs from heterogeneous terrain data and feature databases. We achieved this by refining and implementing the I/O-efficient algorithm by Agarwal *et al.* [1] for computing constrained Delaunay triangulations (CDTs). A CDT is an extension of the regular Delaunay triangulation (DT) with support for enforcing the presence of certain constraint edges in the triangulation.

Our starting point was an existing implementation of DT that did not support constraints. We accomplished the following at the end of Phase I.

1. We extended the DT algorithm to construct a CDT that is efficient when data sets are small enough to fit in main memory.



(a) (b) (c)

Figure 1. (a) (b) An example terrain depicted with contours through saddle vertices and showing the critical vertices of the terrain. (c) The contour tree of the terrain in (a).

2. We implemented the Agarwal *et al.* algorithm, which needs the above algorithm as a subroutine. Their algorithm assumes that the number of constraints are small enough to fit in the main memory. We made some progress on relaxing this constraint, and the work in Phase I has suggested idea how to relax this assumption completely.

The work in Phase I on CDT demonstrated the feasibility of constructing a scalable algorithm. It also suggested that significant work is often needed even after enforcing constraints in the construction of a DEM. Constraints are represented as polygonal chains in 2D and it requires work to appropriately embed them in 3D. Furthermore, polygonal chains consists of line segments but the features they represent often have a width (e.g. highways, rivers) or an interior (e.g. houses and lakes). Inserting the line segment itself as a constraint is thus not sufficient to ensure that roads are flat and lakes and rooftops are appropriately represented in the presence of interiour LiDAR data. Part of our Phase II effort will be to deal with this problem.

Dynamic contour trees. Although not originally proposed in the Phase I proposal, while working on DEM construction we realized that auxiliary structures are needed to perform terrain analysis efficiently. One such auxiliary structure, which is useful in many applications, such as denoising, hydrology analysis, and contour maps, is the so-called *contour tree* [3]. Roughly speaking, the nodes of a contour tree are the critical points (minima, maxima, and saddle points) and its edges represent the evolution of contours (i.e., connected components of level sets of a terrain) as the height changes – when a new contour appears, when an existing contour disappears, when two contours merge, or when a contour splits into two. See Figure 1 for an example.

Therefore, if the DEM is dynamically updated, then the contour tree needs to be updated as well. We thus developed a simple, efficient algorithm for maintaining the contour tree of a terrain, as the DEM is updated—either heights at certain points change, new points are added, or existing points are deleted. The algorithm transforms each such update into a continuous deformation process. The contour tree does not change during this process except when certain *critical events* occur. We characterize the changes at each event, and show that each event causes simple changes in the tree. We are currently writing this result and will be submitting it for publication later this summer.

C2 - Denoising. Most terrain flow-analysis algorithms assume water flows downhill until it reaches a local minimum (or sink). Several anthropogenic features (e.g. bridges) obstruct the water flow and the denoising component corrects many of these issues algorithmically. In Phase I we developed an algorithm to detect bridges in the DEM located in the vicinity of sinks, because they have major effects on the analysis, see Figure 4 for an example.

The algorithm divides the terrain into *watersheds*; one for each sink. The watershed of a sink is the area of the terrain that drains to the sink. We observed that bridges often create artificial sinks on watershed boundaries (e.g. a river flows downstream until it hits a sink right next to the bridge). Thus if a bridge is the cause of an artificial sink in the DEM, the sink will be very close to the bridge itself and the terrain across the bridge will likely be associated with a different watershed. Based on this observation we developed a simple algorithm to identify potential bridges.

This effort demonstrates that it is feasible to detect bridges that are blocking significant hidden flow paths, but manual validation is needed as the simple geometric approach generates a few false positives. We have implemented preliminary support for integrating the results in the online service, mentioned later. This will allow the user to quickly scan an area and mark incorrectly identified bridges.

C3 - Analysis-driven level of detail. Originally we had proposed to investigate visibility-preserving levels of detail (LoD) algorithm for a terrain. In addition to visibility-preserving terrain LOD, we also studied the problem of computing a hierarchical DEM that preserves the lengths of shortest paths on the terrain.

A commonly used approach to build a hierarchical DEM from a triangulation is the so-called *edge-contraction* algorithm [4], which at each stage contracts an edge to a single vertex; see Figure 2. We developed an edge-contraction algorithm that estimates how the contraction of an edge degrades the quality of a path on the terrain, computes the *optimal* location of the contracted point, and chooses an edge for contraction that causes minimum degradation in the quality of paths. In particular, we designed a *penalty function* π, so that $\pi(e)$ for edge e estimates the average distortion of shortest paths on the terrain if e is contracted to a single point. We developed an efficient algorithm for updating the penalty function when an edge is contracted to a single point. At each step, the edge with the smallest penalty is contracted, and the penalties of the affected edges are updated. Our experiments show that our approach performs significantly better than the previous ones and is quite efficient. Details can be found in the paper [5], which has been submitted to *International Journal of Robotics Research* for publication.

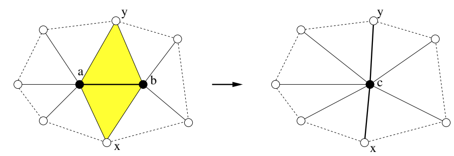
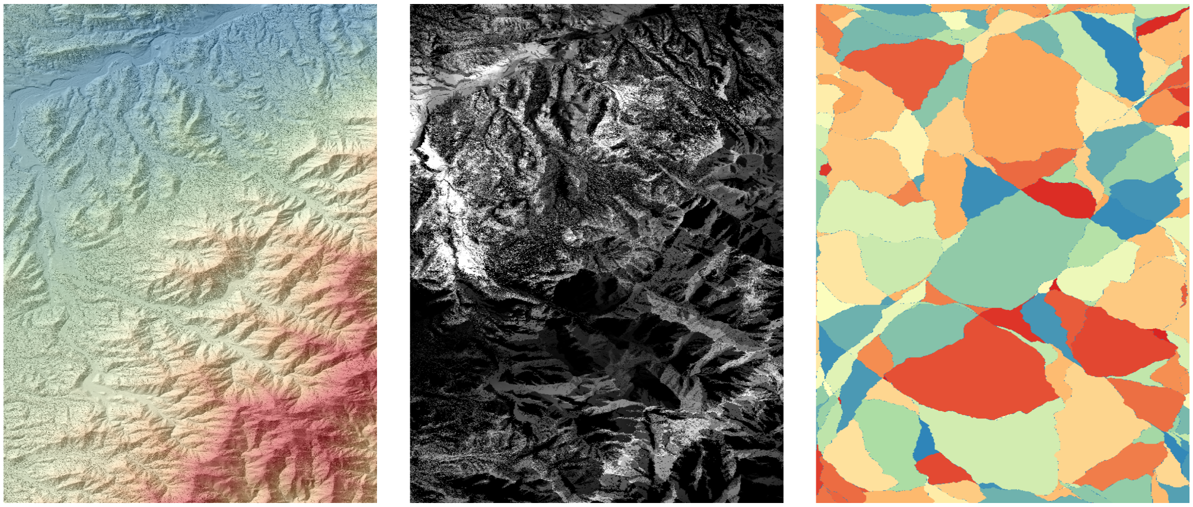


Figure 2. Edge contraction.

Since visibility is a very broad concept, we focused on visibility with the goal of computing highly occluded paths on terrains, a problem of enormous interest in army applications. We faced two challenges developing an edge-contraction algorithm for this case. First, unlike shortest paths, it seems difficult to develop a simple penalty function that estimates how much visibility information is compromised by contracting a single edge. Second, the size of triangles varied significantly on the simplified terrain, which made it difficult to compute highly occluded paths on simplified terrains. We therefore pursued a different approach, which addressed both of these challenges.

We hypothesized that a sparse one-dimensional network on the terrain can be constructed that partitions the terrain into a small number of regions so that the highly occluded path either follows a shortest path inside each region or follows the boundary of these regions; see Figure 3. This not only reduces the size of the terrain but also simplifies the path-computation algorithm. We developed a simple learning-based algorithm to verify our hypothesis. Our experiments show that one can indeed construct a sparse 1D network to find highly occluded paths on a terrain, but the algorithm for constructing the network is slow. The next step is to design an algorithm that exploits the geometry and topology of the terrain.

C4 - Uncertainty aware algorithms. We designed an out-of-core algorithm to construct a stochastic representation of a grid DEM from LiDAR data, which models the uncertainty in the original LiDAR data because of measurement errors. This algorithm first constructs a hierarchical partition of the region for which we wish to build the DEM, next, for each cell of the resulting partition it chooses LiDAR points inside the cell and its neighborhood, models the terrain inside the cell as a stochastic process (e.g. a Gaussian process) [2], and then uses this model to infer the elevation at each grid point inside the cell. Although the second step is computationally expensive, the hierarchical partition ensures that the number of points used to compute the



(a) (b) (c)

Figure 3. Building a 1D sparse network. (a) DEM of the terrain; (b) visibility map from a given set of guards; (c) partition of the terrain into regions whose boundaries form the 1D network.

Gaussian process is small. This is just the first step and significant research will be conducted on Phase II.

C5 - Online service. The overall objective of the online service is to integrate the advantages of analysis driven modeling with users and help transcend the notion of a single DEM that can be applied to all tasks at all times. Our Phase I objective was to build a prototype version of the service. We have developed a scalable and fully-functional system for visualizing national DEMs and a wealth of derived products on these DEMs. We recently launched SCALGO Live Global[1] in collaboration with SCALGO DK[2], which allows users to visualize, and interact with, a near-global terrain model[3] and a set of derived hydrological products. The launch was a success and was featured in major online venues (e.g. Slashdot); we served 18.3 million map tiles over the span of a few days. We have since added the full 10m NED (along with some derived products) to the online service[4].

On top of this system we developed prototypes of some of the components supporting our vision of providing custom, high-quality APDEMs in a dynamic and interactive setting: (1) Support for accessing and downloading the underlying DEM for an area of interest. (2) Submitting, editing and vizualising user-supplied corrections to the DEM in the form of polygonal chains and using on-the-fly queries to our back-end to retrieve the elevation at the chain vertices, refer to Figure 4. (3) Simulated support for dynamically updating the model using the submitted corrections, and using periodic re-computations though an automated system for managing the APDEM construction process across a set of computing machines - an important part of managing large amounts of data and APDEMs with minimal manual labor.

The speed of the data service is important to ensure users enjoy working with the prototype. Our data service prototype has been carefully implemented to be multi-threaded, heavily cached and disk-efficient - this ensures that the data service itself is not the bottleneck in the system. In practice, network latency as the primary bottleneck for users that are not geographically close to our servers. Due to both user and server-side caching, this latency is not detrimental to the experience, but a decrease in latency would result in a better user experience for such users.

[1] Freely available at `http://scalgo.com/live/global`.

[2] Denmark-based company by the same group of founders focusing on terrain analysis.

[3] We used the 3 arc seconds (90m at the equator) model produced by the The Shuttle Radar Topography Mission (SRTM)

[4] Available at `http://scalgo.com/live/ned`.

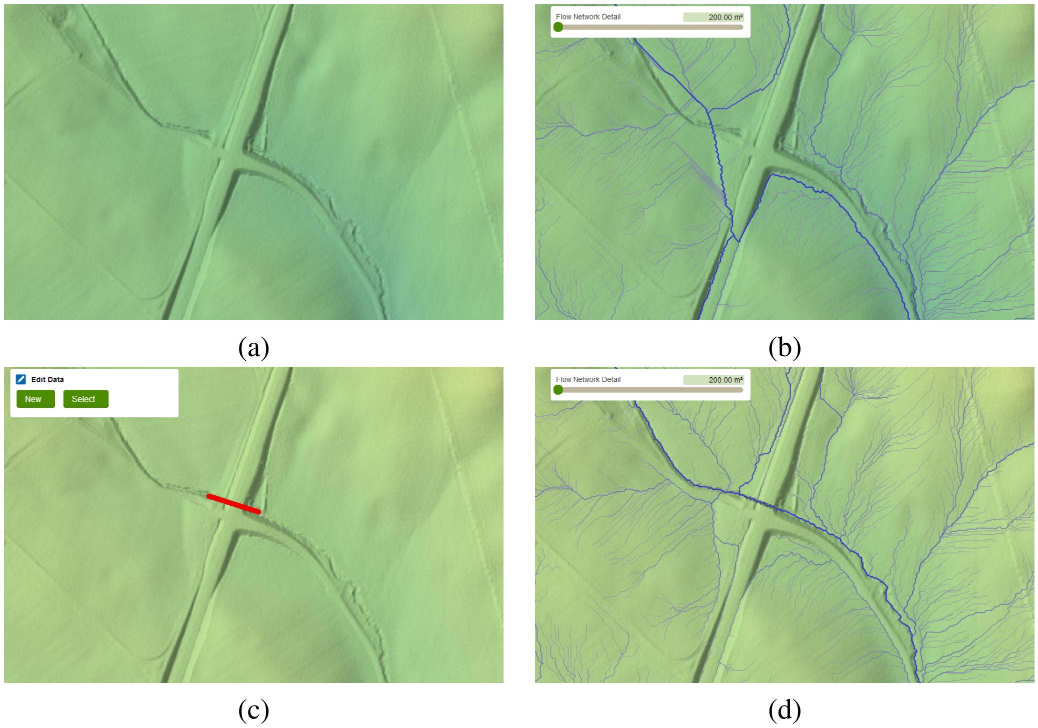


Figure 4. Screenshots from the online service showing (a) a ∼5ft (1.6m) DEM zoomed in on an highway bridge spanning a stream. (b) River network at this part of the DEM, water flows from top left to bottom right. The stream is incorrectly diverged at the highway bridge and is mapped to the wrong side of the street after the bridge. (c) An edit is inserted, cutting through the highway at the bridge, clearing room for the stream. (d) Correct mapping of stream.

Phase I Conclusion. Overall Phase I was very successful and clearly demonstrates the feasilibity of our overall vision, both in terms of techonlogy and commercialization. Figure 5 summarizes our progress towards the Phase I objectives for each of the five components.

- **(C1)** We produced the CDT algorithm and learned that more work will be needed to support very large numbers of constraints. We developed an algorithm for dynamic contour trees, an addition to the content in the Phase I proposal.
- **(C2)** We tested a geometric approach for detecting bridges obstructing major flow paths; more work is needed to increase the performance and accuracy.
- **(C3)** We investigated how to approximate the visibility map for computing occluded paths, and how to simplify the terrain while maintaining shortest paths.
- **(C4)** We designed an algorithm for generating a stochastic DEM. More work is needed on the effectiveness of the algorithm and on basing the uncertainty information on the full LiDAR waveform of the input points.
- **(C5)** We developed a fully functional system for launching SCALGO Live Global and the prototype functionality required for the update and download capabilities was functional as well.

Besides the experience in project feasibility gained from Phase I we produced a number of tangible outcomes. Some of these are mentioned above, but we summarize them here. We have prepared a publication on the maintenance of dynamic contour trees and plan to submit this for publication in the coming months. We submitted a paper [5] on level of detail using edge contractions for publication to the *International*

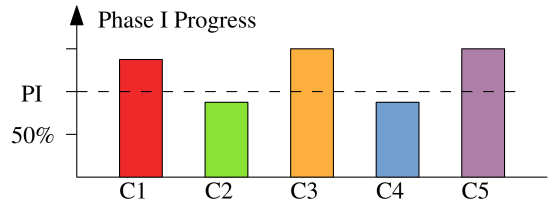


Figure 5. Overall progress on the tasks during Phase 1 as a percent of the proposed objectives. The dashed line indicates the work proposed in the Phase I proposal.

Journal of Robotics Research. We are also in the process of writing a paper on extracting networks for finding low-visibility paths, this will be submitted in the comings months as well.

Finally, we acquired the full 10m National Elevation Dataset (NED) model for the contiguous US states produced by the USGS. This model, periodically updated by the USGS, is an excellent starting point for a good national model. The model consists of about 200 billion cells and is now available in its entirety on our online service. We have also produced derived products on the entire model and we are a first mover in our ability to perform these computation on such a large model. Besides being interesting in their own right, these computations are important drivers for identifying missing flow paths in the model, as as baselines for comparisons with future models.

References

[1] P. K. Agarwal, L. Arge, and K. Yi. I/O-efficient construction of constrained Delaunay triangulations. In *Proc. European Sympos. Algorithms*, pages 355–366, 2005.

[2] S. Banerjee, B. Carlin, and A. E. Gelfand. *Hierarchical Modeling and Analysis for Sptial Data*. Chapman and Hall, New York, 2004.

[3] H. Carr, J. Snoeyink, and U. Axen. Computing contour trees in all dimensions. *Comput. Geom.*, 24(2):75–94, 2003.

[4] M. Garland and P. S. Heckbert. Surface simplification using quadric error metrics. In *SIGGRAPH*, pages 209–216, 1997.

[5] D. Shaharabani, O. Salzman, P. K. Agarwal, and D. Halperin. Sparsification of motion-planning roadmaps by edge contraction. submitted to Intl. J. Robotics Research.